#34
Entered 5/3/01

1653

RAW SEQUENCE LISTING DATE: 05/03/2001
PATENT APPLICATION: US/09/105,117I TIME: 12:46:45

ENTERED

Input Set : A:\LysG_orf3.txt
Output Set: N:\CRF3\05032001\I105117I.raw

```
     5 <110> APPLICANT: Forschungszentrum Juelich GmbH; Marina Vrlijc et al.
     9 <120> TITLE OF INVENTION: Process for the microbial production of amino acids by
    11       boosted activity of export carriers
    15 <130> FILE REFERENCE: 1
C--> 19 <140> CURRENT APPLICATION NUMBER: US/09/105,117I
C--> 21 <141> CURRENT FILING DATE: 1998-06-17
    25 <160> NUMBER OF SEQ ID NOS: 3
    29 <170> SOFTWARE: PatentIn Ver. 2.0
    33 <210> SEQ ID NO: 1
    35 <211> LENGTH: 2374
    37 <212> TYPE: DNA
    39 <213> ORGANISM: Corynebacterium glutamicum
    43 <220> FEATURE:
W--> 45 <221> NAME/KEY: CDS ORF3 (partial)
    47 <222> LOCATION: (2)..(652)
    51 <220> FEATURE:
W--> 53 <221> NAME/KEY: CDS LysG (Regulator lysE)
    55 <222> LOCATION: (1421)..(2293)
    59 <400> SEQUENCE: 1
    61 a gat act cct ttg gaa gaa acc atg tac gca ttg cgt gac att gtt gcg 49
    63   Asp Thr Pro Leu Glu Glu Thr Met Tyr Ala Leu Arg Asp Ile Val Ala
    65     1               5                  10                  15
    69 tct gga aag gct ctt tac gtg ggt att tct tcc tac ggt cca gag ctc   97
    71 Ser Gly Lys Ala Leu Tyr Val Gly Ile Ser Ser Tyr Gly Pro Glu Leu
    73              20                  25                  30
    77 aca gcg gag gcg gct gag ttc atg gcg gag gag ggc tgc ccg ctt ctg   145
    79 Thr Ala Glu Ala Ala Glu Phe Met Ala Glu Glu Gly Cys Pro Leu Leu
    81          35                  40                  45
    85 att cat cag cca agc tat tcc atc att aat cgt tgg gtg gag gaa ccg   193
    87 Ile His Gln Pro Ser Tyr Ser Ile Ile Asn Arg Trp Val Glu Glu Pro
    89      50                  55                  60
    93 ggc gat gac ggt gag aac ttg ttg cag tca gct gcc aac aat ggt ctt   241
    95 Gly Asp Asp Gly Glu Asn Leu Leu Gln Ser Ala Ala Asn Asn Gly Leu
    97  65                  70                  75                  80
   101 ggc gtc att gct ttc tca cca ctt gcg cag ggc ctg ctc acg gac aaa   289
   103 Gly Val Ile Ala Phe Ser Pro Leu Ala Gln Gly Leu Leu Thr Asp Lys
   105                 85                  90                  95
   109 tat ctc gat gga att cca gag ggt tcc cgc gcc agc cag ggt aag tcc   337
   111 Tyr Leu Asp Gly Ile Pro Glu Gly Ser Arg Ala Ser Gln Gly Lys Ser
   113             100                 105                 110
   117 ctg tct gag ggc atg ttg aac gtg aac aat att gat atg gtc cgc aag   385
   119 Leu Ser Glu Gly Met Leu Asn Val Asn Asn Ile Asp Met Val Arg Lys
   121         115                 120                 125
   125 ctc aat gac atc gcc cag gaa cgc ggg cag tca ctt gcg cag atg gcg   433
   127 Leu Asn Asp Ile Ala Gln Glu Arg Gly Gln Ser Leu Ala Gln Met Ala
   129     130                 135                 140
   133 ctt gca tgg gtg ctg cgc gag caa gga gag tac ggc gcg gat acc gtg   481
```

RAW SEQUENCE LISTING DATE: 05/03/2001
PATENT APPLICATION: US/09/105,117I TIME: 12:46:45

Input Set : A:\LysG_orf3.txt
Output Set: N:\CRF3\05032001\I105117I.raw

```
     135 Leu Ala Trp Val Leu Arg Glu Gln Gly Glu Tyr Gly Ala Asp Thr Val
     137 145                 150                 155                 160
     141 acc agt gca ttg att ggt gct tcg tca gtt gag cag ctg gac aac agc   529
     143 Thr Ser Ala Leu Ile Gly Ala Ser Ser Val Glu Gln Leu Asp Asn Ser
     145                 165                 170                 175
     149 ctt gat tca ctc aac aac ttg gag ttt tct gac gcc gag ttg gag gcg   577
     151 Leu Asp Ser Leu Asn Asn Leu Glu Phe Ser Asp Ala Glu Leu Glu Ala
     153             180                 185                 190
     157 atc gat gag att tcc cac gac gcc ggc atc aac att tgg gcg aag gcc   625
     159 Ile Asp Glu Ile Ser His Asp Ala Gly Ile Asn Ile Trp Ala Lys Ala
     161         195                 200                 205
     165 acc gat tcc aaa acc cgc gaa aac taa cccatcaaca tcagtttgat         672
     167 Thr Asp Ser Lys Thr Arg Glu Asn
     169     210                 215
     173 ggccaatgcg gtcatcacaa ctgccacgac gacgttgatc cagcgccaca ccttggggct 732
     177 ggacagcggg cgtgacaatg ctgctgcgcc gaaacccacc agcgggaacc agatcaggct 792
     181 tgccgcgaac gcgccagcgg cgaaaatcca ccgtccggtg tcgccgtatt gcgcgccgac 852
     185 gccgccgata aacacaaacg cgtccaaata cgcattcggg ttcaaccagg tcagcacgat 912
     189 tgccatcaac atgggcttta cccaaacccg ctgcttatcg acgctcacct ccacccgcac 972
     193 ccggttgcgc gtgtcagtgg ccaccgccga accgcccaaa ggcgtgtcat cgggcacggt 1032
     197 tggttctgtt tcttcaatga tctgtggcgc ttccaccttg tttgtcatgg cgtctttcgc 1092
     201 tgccatgacg gcaaaccata acaggtaagc gatgccaccc cagcgcataa tatcgagcac 1152
     205 gatcggcgcg gcattggaca aaagatcaac gcccaaggtg ccggcgatga acaaaaagac 1212
     209 gtcagaaatt aaacacacga gaagaaccgc aatgagtcct tcgcgcttaa ttccttgttt 1272
     213 aatcaccagt acattctgcg gtccgatgga cagtaaaaga ctggcccccа aaagcagacc 1332
     217 tgtaatgaag atttccatga tcaccatcgt gacctatgga agtacttaag taaaatgatt 1392
     221 ggttcttaac atggtttaat atagcttc atg aac ccc att caa ctg gac act    1444
     223                            Met Asn Pro Ile Gln Leu Asp Thr
W--> 225                                    220                 225
     229 ttg ctc tca atc att gat gaa ggc agc ttc gaa ggc gcc tcc tta gcc   1492
     231 Leu Leu Ser Ile Ile Asp Glu Gly Ser Phe Glu Gly Ala Ser Leu Ala
W--> 233                 230                 235                 240
     237 ctt tcc att tcc ccc tcg gcg gtg agt cag cgc gtt aaa gct ctc gag   1540
     239 Leu Ser Ile Ser Pro Ser Ala Val Ser Gln Arg Val Lys Ala Leu Glu
W--> 241             245                 250                 255
     245 cat cac gtg ggt cga gtg ttg gta tcg cgc acc caa ccg gcc aaa gca   1588
     247 His His Val Gly Arg Val Leu Val Ser Arg Thr Gln Pro Ala Lys Ala
W--> 249         260                 265                 270
     253 acc gaa gcg ggt gaa gtc ctt gtg caa gca gcg cgg aaa atg gtg ttg   1636
     255 Thr Glu Ala Gly Glu Val Leu Val Gln Ala Ala Arg Lys Met Val Leu
W--> 257     275                 280                 285
     261 ctg caa gca gaa act aaa gcg caa cta tct gga cgc ctt gct gaa atc   1684
     263 Leu Gln Ala Glu Thr Lys Ala Gln Leu Ser Gly Arg Leu Ala Glu Ile
W--> 265 290                 295                 300                 305
     269 ccg tta acc atc gcc atc aac gca gat tcg cta tcc aca tgg ttt cct   1732
     271 Pro Leu Thr Ile Ala Ile Asn Ala Asp Ser Leu Ser Thr Trp Phe Pro
W--> 273                 310                 315                 320
     277 ccc gtg ttc aac gag gta gct tct tgg ggt gga gca acg ctc acg ctg   1780
     279 Pro Val Phe Asn Glu Val Ala Ser Trp Gly Gly Ala Thr Leu Thr Leu
```

RAW SEQUENCE LISTING DATE: 05/03/2001
PATENT APPLICATION: US/09/105,117I TIME: 12:46:45

Input Set : A:\LysG_orf3.txt
Output Set: N:\CRF3\05032001\I105117I.raw

```
W--> 281              325                 330                 335
     285 cgc ttg gaa gat gaa gcg cac aca tta tcc ttg ctg cgg cgt gga gat   1828
     287 Arg Leu Glu Asp Glu Ala His Thr Leu Ser Leu Leu Arg Arg Gly Asp
W--> 289         340                 345                 350
     293 gtt tta gga gcg gta acc cgt gaa gct aat ccc gtg gcg gga tgt gaa   1876
     295 Val Leu Gly Ala Val Thr Arg Glu Ala Asn Pro Val Ala Gly Cys Glu
W--> 297     355                 360                 365
     301 gta gta gaa ctt gga acc atg cgc cac ttg gcc att gca acc ccc tca   1924
     303 Val Val Glu Leu Gly Thr Met Arg His Leu Ala Ile Ala Thr Pro Ser
W--> 305 370                 375                 380                 385
     309 ttg cgg gat gcc tac atg gtt gat ggg aaa cta gat tgg gct gcg atg   1972
     311 Leu Arg Asp Ala Tyr Met Val Asp Gly Lys Leu Asp Trp Ala Ala Met
W--> 313                 390                 395                 400
     317 ccc gtc tta cgc ttc ggt ccc aaa gat gtg ctt caa gac cgt gac ctg   2020
     319 Pro Val Leu Arg Phe Gly Pro Lys Asp Val Leu Gln Asp Arg Asp Leu
W--> 321              405                 410                 415
     325 gac ggg cgc gtc gat ggt cct gtg ggg cgc agg cgc gta tcc att gtc   2068
     327 Asp Gly Arg Val Asp Gly Pro Val Gly Arg Arg Arg Val Ser Ile Val
W--> 329         420                 425                 430
     333 ccg tcg gcg gaa ggt ttt ggt gag gca att cgc cga ggc ctt ggt tgg   2116
     335 Pro Ser Ala Glu Gly Phe Gly Glu Ala Ile Arg Arg Gly Leu Gly Trp
W--> 337     435                 440                 445
     341 gga ctt ctt ccc gaa acc caa gct gct ccc atg cta aaa gca gga gaa   2164
     343 Gly Leu Leu Pro Glu Thr Gln Ala Ala Pro Met Leu Lys Ala Gly Glu
W--> 345 450                 455                 460                 465
     349 gtg atc ctc ctc gat gag ata ccc att gac aca ccg atg tat tgg caa   2212
     351 Val Ile Leu Leu Asp Glu Ile Pro Ile Asp Thr Pro Met Tyr Trp Gln
W--> 353                 470                 475                 480
     357 cga tgg cgc ctg gaa tct aga tct cta gct aga ctc aca gac gcc gtc   2260
     359 Arg Trp Arg Leu Glu Ser Arg Ser Leu Ala Arg Leu Thr Asp Ala Val
W--> 361              485                 490                 495
     365 gtt gat gca gca atc gag gga ttg cgg cct tag ttacttctga aaaggttcag 2313
     367 Val Asp Ala Ala Ile Glu Gly Leu Arg Pro
W--> 369         500                 505
     373 ggtttttcac ttcttcgccc gcaggaattg ggccaggcag agtaacacct tcagcaaatg 2373
     377 g                                                               2374
     383 <210> SEQ ID NO: 2
     385 <211> LENGTH: 216
     387 <212> TYPE: PRT ORF3 (partial)
     389 <213> ORGANISM: Corynebacterium glutamicum
     393 <400> SEQUENCE: 2
     395 Asp Thr Pro Leu Glu Glu Thr Met Tyr Ala Leu Arg Asp Ile Val Ala
     397   1               5                  10                  15
     401 Ser Gly Lys Ala Leu Tyr Val Gly Ile Ser Ser Tyr Gly Pro Glu Leu
     403             20                  25                  30
     407 Thr Ala Glu Ala Ala Glu Phe Met Ala Glu Glu Gly Cys Pro Leu Leu
     409         35                  40                  45
     413 Ile His Gln Pro Ser Tyr Ser Ile Ile Asn Arg Trp Val Glu Glu Pro
     415     50                  55                  60
```

RAW SEQUENCE LISTING DATE: 05/03/2001
PATENT APPLICATION: US/09/105,117I TIME: 12:46:45

Input Set : A:\LysG_orf3.txt
Output Set: N:\CRF3\05032001\I105117I.raw

```
419 Gly Asp Asp Gly Glu Asn Leu Leu Gln Ser Ala Ala Asn Asn Gly Leu
421  65                  70                  75                  80
425 Gly Val Ile Ala Phe Ser Pro Leu Ala Gln Gly Leu Leu Thr Asp Lys
427                  85                  90                  95
431 Tyr Leu Asp Gly Ile Pro Glu Gly Ser Arg Ala Ser Gln Gly Lys Ser
433             100                 105                 110
437 Leu Ser Glu Gly Met Leu Asn Val Asn Asn Ile Asp Met Val Arg Lys
439         115                 120                 125
443 Leu Asn Asp Ile Ala Gln Glu Arg Gly Gln Ser Leu Ala Gln Met Ala
445     130                 135                 140
449 Leu Ala Trp Val Leu Arg Glu Gln Gly Glu Tyr Gly Ala Asp Thr Val
451 145                 150                 155                 160
455 Thr Ser Ala Leu Ile Gly Ala Ser Ser Val Glu Gln Leu Asp Asn Ser
457                 165                 170                 175
461 Leu Asp Ser Leu Asn Asn Leu Glu Phe Ser Asp Ala Glu Leu Glu Ala
463             180                 185                 190
467 Ile Asp Glu Ile Ser His Asp Ala Gly Ile Asn Ile Trp Ala Lys Ala
469         195                 200                 205
473 Thr Asp Ser Lys Thr Arg Glu Asn
475     210                 215
481 <210> SEQ ID NO: 3
483 <211> LENGTH: 290
485 <212> TYPE: PRT LysG (Regulator lysE)
487 <213> ORGANISM: Corynebacterium glutamicum
491 <400> SEQUENCE: 3
493 Met Asn Pro Ile Gln Leu Asp Thr Leu Leu Ser Ile Ile Asp Glu Gly
495   1               5                  10                  15
499 Ser Phe Glu Gly Ala Ser Leu Ala Leu Ser Ile Ser Pro Ser Ala Val
501              20                  25                  30
505 Ser Gln Arg Val Lys Ala Leu Glu His His Val Gly Arg Val Leu Val
507          35                  40                  45
511 Ser Arg Thr Gln Pro Ala Lys Ala Thr Glu Ala Gly Glu Val Leu Val
513      50                  55                  60
517 Gln Ala Ala Arg Lys Met Val Leu Leu Gln Ala Glu Thr Lys Ala Gln
519  65                  70                  75                  80
523 Leu Ser Gly Arg Leu Ala Glu Ile Pro Leu Thr Ile Ala Ile Asn Ala
525                  85                  90                  95
529 Asp Ser Leu Ser Thr Trp Phe Pro Pro Val Phe Asn Glu Val Ala Ser
531             100                 105                 110
535 Trp Gly Gly Ala Thr Leu Thr Leu Arg Leu Glu Asp Glu Ala His Thr
537         115                 120                 125
541 Leu Ser Leu Leu Arg Arg Gly Asp Val Leu Gly Ala Val Thr Arg Glu
543     130                 135                 140
547 Ala Asn Pro Val Ala Gly Cys Glu Val Val Glu Leu Gly Thr Met Arg
549 145                 150                 155                 160
553 His Leu Ala Ile Ala Thr Pro Ser Leu Arg Asp Ala Tyr Met Val Asp
555                 165                 170                 175
559 Gly Lys Leu Asp Trp Ala Ala Met Pro Val Leu Arg Phe Gly Pro Lys
561             180                 185                 190
```

RAW SEQUENCE LISTING DATE: 05/03/2001
PATENT APPLICATION: US/09/105,117I TIME: 12:46:45

Input Set : A:\LysG_orf3.txt
Output Set: N:\CRF3\05032001\I105117I.raw

```
565 Asp Val Leu Gln Asp Arg Asp Leu Asp Gly Arg Val Asp Gly Pro Val
567         195                 200                 205
571 Gly Arg Arg Arg Val Ser Ile Val Pro Ser Ala Glu Gly Phe Gly Glu
573     210                 215                 220
577 Ala Ile Arg Arg Gly Leu Gly Trp Gly Leu Leu Pro Glu Thr Gln Ala
579 225                 230                 235                 240
583 Ala Pro Met Leu Lys Ala Gly Glu Val Ile Leu Leu Asp Glu Ile Pro
585                 245                 250                 255
589 Ile Asp Thr Pro Met Tyr Trp Gln Arg Trp Arg Leu Glu Ser Arg Ser
591             260                 265                 270
595 Leu Ala Arg Leu Thr Asp Ala Val Val Asp Ala Ala Ile Glu Gly Leu
597         275                 280                 285
601 Arg Pro
603     290
```

VERIFICATION SUMMARY DATE: 05/03/2001
PATENT APPLICATION: US/09/105,117I TIME: 12:46:46

Input Set : A:\LysG_orf3.txt
Output Set: N:\CRF3\05032001\I105117I.raw

L:19 M:270 C: Current Application Number differs, Replaced Current Application Number
L:21 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:45 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:53 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:225 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:233 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:241 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:249 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:257 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:265 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:273 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:281 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:289 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:297 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:305 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:313 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:321 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:329 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:337 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:345 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:353 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:361 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:369 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1